

Metadata at the Federal Reserve

Linda Powell

XBRL W3 Workshop 2009

The views expressed in this paper are those of the author and do not necessarily represent those of the Board of Governors of the Federal Reserve System.



Background

- Board of Governors of the Federal Reserve System
 - U.S. central bank – engages in monetary policy and banking supervision
- Improvements in technology led to improved data availability (more and better)
- Changes in the financial markets resulted in the need for more data



Historically

- Purchased Data were used for specific research projects
 - Therefore, economists that purchased data maintained the data themselves
- Data collected by the Federal Reserve System were used for policy work
 - Therefore, data maintenance was centralized and well documented



Business Problem

Currently no matter how data enter the Federal Reserve System (FRS) they may be used for policy work as well as research projects

Data collected by the FRS was well documented and centralized

Data purchased was decentralized and not centrally documented



As of October 2009

The Board has three metadata repositories

1. FRS collected data (MDRM)
2. Collection level metadata (DANCE)
3. Vendor metadata repository

All three serve a purpose but they didn't work together




MDRM

- Collection level metadata
- Variable level metadata
 - Prefix identifies the data collection
 - E.g. bank financials vs. finance company financials
 - Suffix identifies the accounting concept
 - Uniform across data collections
 - Total assets always has the suffix 2170
 - When you combine the prefix and suffix you get the exact concept
 - RCON2170 is total assets for a bank
 - DFCR2170 is total assets for a finance company
- Covers financial, structure, and supervisory data



MDRM Demo

<http://www.federalreserve.gov/>



Users want more and they want it in one retrieval

The MDRM has been around for more than 30 years. It contains collection and variable level metadata.

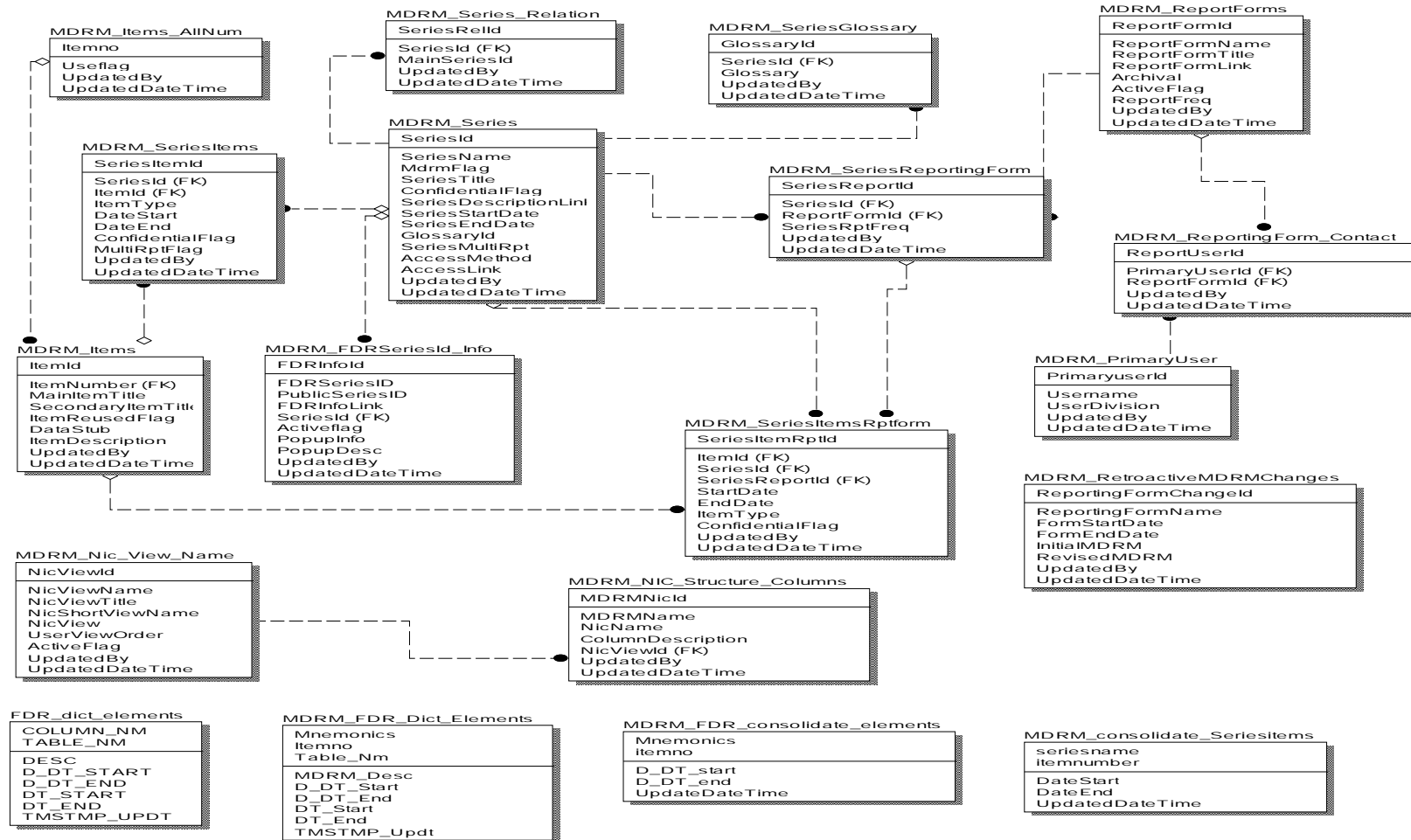
As soon as DANCE was implemented in 2004 users wanted more:

- Lots more metadata
- A way to make centralized notes about datasets

It was time to look at international standards

So we decided to look at enhancing all three metadata repositories

MDRM Schema



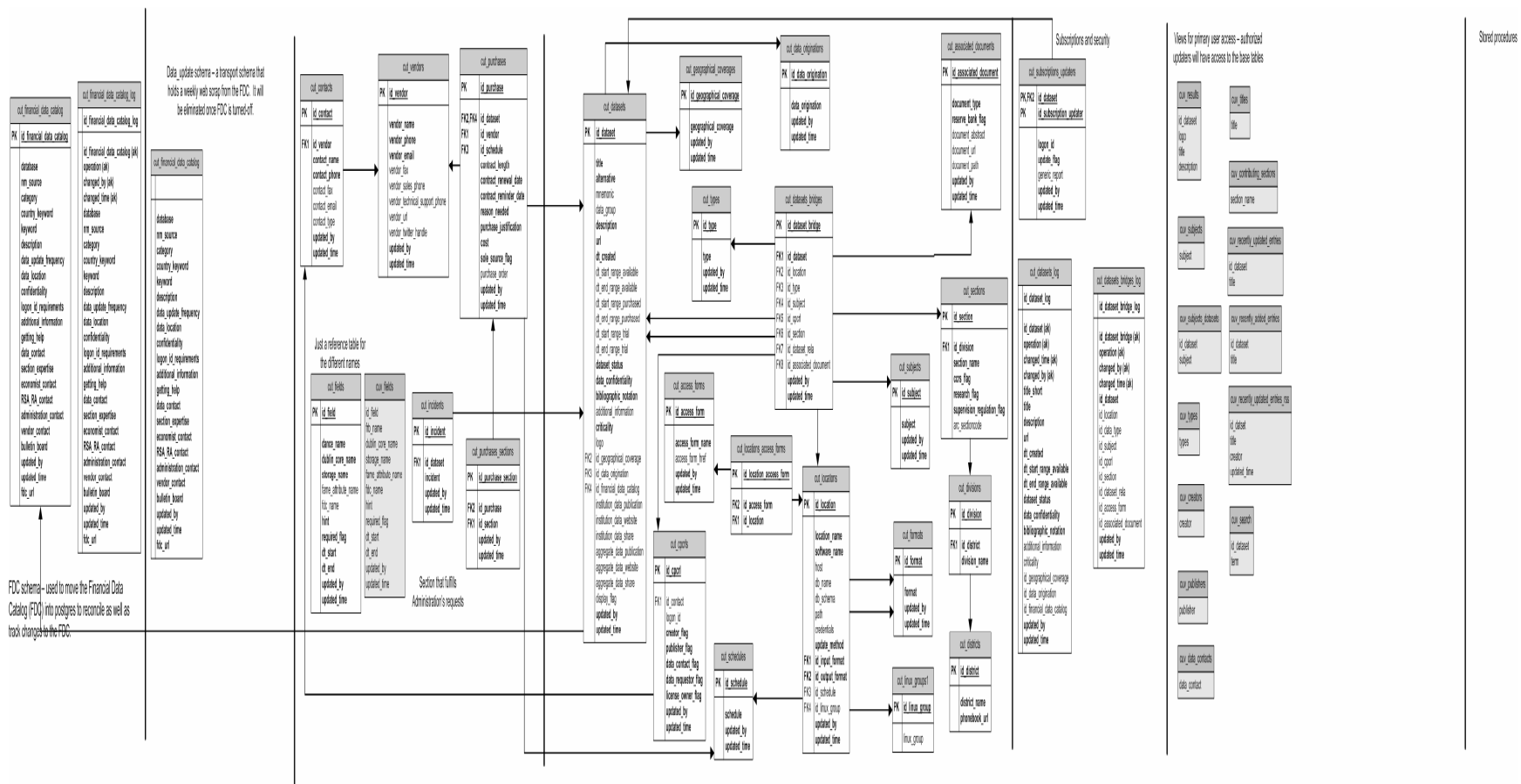
Dance Variables

Original vs. Dublin Core

Original Variables	FRB Name of Retained/ New Variables	Dublin Core Element
category		
database	title	title
	alternative	alternative
database URL	product url	
description	description	description
keyword(s)	keyword	subject
vendor	publisher	publisher
vendor URL	vendor url	
	creator	creator
	record id	identifier
	date created	date
	date range available	available
	geographical coverage	coverage
status	dataset status	accrual policy
	data origination	source
	related resources	relation

Original Variables	FRB Name of Retained/ New Variables	Dublin Core Element
form of access	data location	
form of access	format	format
	type	type
	physical medium	medium
	update method	accrual method
	software name	software
	update schedule	accrual periodicity
Data Contact	data contact	
Division	purchasing division	
	data requestor	
	contributing section	contributor
	purchasing section	
	vendor contacts	
license contact	license owner	rights holder
license information	data confidentiality	rights
	bibliographic notation	bibliographic citation
	license agreement	license
	help files	instructional method

DANCE Schema



FDC schema - used to move the Financial Data Catalog (FDC) into postgres to reconcile as well as track changes to the FDC.

Data update schema - a transport schema that holds a weekly web scrap from the FDC. It will be eliminated once FDC is turned-off.

Just a reference table for the different names

Section that fulfills Administrator's requests

Subscriptions and security

Views for primary user access - authorized updaters will have access to the base tables

Stored procedures

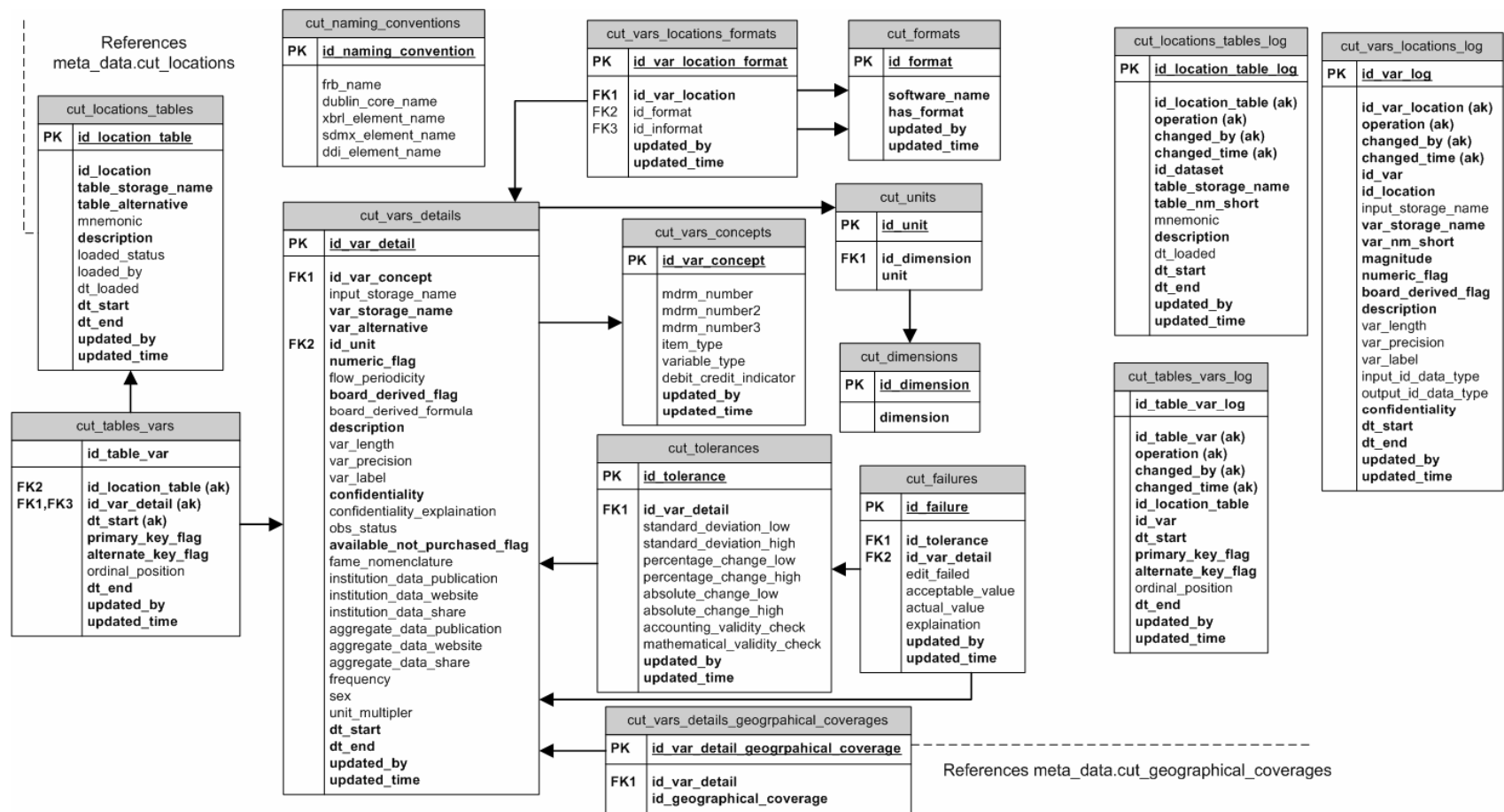
Vendor Metadata Repository

DC and XBRL Variables

FRB Name of Retained/ New Variables	Dublin Core Element	XBRL Element
variable storage name	title	
variable name short	alternative	
mnemonic		
description	description	
mdrm number		
id variable	identifier	
date start		
date end		
ordinal position		
flow-stock		flow-stock
flow periodicity		flow periodicity
debit / credit indicator		debit / credit indicator
data type		data type

FRB Name of Retained/New Variables	Dublin Core Element	XBRL Element
multiplier		
numeric flag		
derived flag		
Formula		formula
variable length		
variable precision		
confidentiality		
format	hasformat	
record last updated		
record updated by		

Vendor Metadata Schema





Conclusion

- Metadata are important for understanding as well as for discovery
- International standards provide breadth and depth
- Having a nomenclature that spans data collections makes understanding of the data easier and reduces the risk of errors by users